PRODUCTS OF RANDOM MATRICES AND q-CATALAN NUMBERS

A.Khorunzhy
UFR de Mathématiques
Université Paris 7 - Denis Diderot, FRANCE*

February 1, 2008

Abstract

We give an interpretation of the q-Catalan numbers in frameworks of the random matrix theory and weighted partitions of the set of integers.

Key words: random matrices, Catalan numbers,
non-commutative random variables
AMS subject classification Primary: 15A52 Secondary: 05A18

In a joint discussion [3], Christian Mazza asked, what one can obtain when regarding the weighted pairings of $2k$ points under the condition that they are non-crossing? In the present note we give one possible answer to this question.

Let us consider a set of N-dimensional random matrices $A^{(r)},\ r \in \mathbf{N}$ determined on the same probability space. We assume that these matrices are real symmetric and

$$[A_N^{(r)}]_{ij} = \frac{1}{\sqrt{N}} a_{ij}^{(r)}, \quad i,j = 1,\ldots,N,$$

where $\{a_{ij}^{(r)}, i \leq j\}$ is the family of Gaussian random variables with zero mathematical expectation and the covariance matrix

$$\mathbf{E}\, a_{ij}^{(r)} a_{i'j'}^{(r')} = V(r-r')\big(\delta_{ii'}\delta_{jj'} + \delta_{ij'}\delta_{i'j}\big). \tag{1}$$

Relations (1) mean that the probability distribution of the matrix $A^{(r)}$ with given r is equal to that of GOE [4]. Certainly we assume V to be real, even, and positively defined function.

*Permanent address: Mathematical Division, Institute for Low Temperature Physics, Kharkov, UKRAINE, e-mail: khorunjy@ilt.kharkov.ua

In this note, our main subject is the product

$$Q_{2k}^{(N)}(p) = \prod_{s=1}^{2k} A_N^{(s)}$$

and we will be related mainly with the simplest case when

$$V(r - r') = p^{|r-r'|}\ ,\ 0 < p < 1. \tag{2}$$

If $N = 1$, we obtain the product of real gaussian random variables. In this case we can use the integration by parts formula for a Gaussian random vector $\vec{\gamma} = (\gamma_1, \dots, \gamma_n)$ with zero mean value:

$$\mathbf{E}\, \gamma_l F(\vec{\gamma}) = \sum_{m=1}^{n} \mathbf{E}\, \gamma_l \gamma_m \mathbf{E}\, \frac{\partial F(\vec{\gamma})}{\partial \gamma_m}, \tag{3}$$

where F is a nonranom function. Then the mathematical expectation

$$\mathbf{E}\, Q_{2k}^{(1)}(p) = \sum_{\omega \in \Omega_{2k}} \prod_{(l,m)\in\omega} p^{|l-m|} \tag{4}$$

is given by the sum of weighted partitions of the set $(1, 2, \dots, 2k)$ of $2k$ labelled points into k pairs. Here ω denotes one particular partition, Ω_{2k} stands for the set of all possible partitions, and $(l, m) \in \omega$ means that given a partition ω, the product is taken over pairs that form ω and whose elements are denoted by (l, m).

Certainly, if $p = 1$, then

$$\mathbf{E}\, Q_{2k}^{(1)}(1) = \frac{(2k)!}{2^k k!}.$$

It is interesting to study (4) in the limit $k \to \infty$, namely, the asymptotic behaviour of the value

$$\frac{1}{k}\ \log\left\{\mathbf{E}\, Q_{2k}^{(N)}(p)\right\},\quad N = 1,\ k \to \infty \tag{5}$$

in the dependence of the parameter p. It is proved that there exists critical value p_c such that (5) is positive for $p > p_c$ and negative for $p < p_c$ [3].

Now let us turn to the case of non-commutative random variables, namely, to the limit of $N = \infty$.

Theorem 1.1
The average value of $Q_{2k}^{(N)}(p)$

$$\frac{1}{N}\ \mathrm{Tr}\left\{A^{(1)} A^{(2)} \cdots A^{(2k)}\right\} = \langle Q_{2k}^{(N)}(p)\rangle \tag{6}$$

converges as $N \to \infty$ in average

$$\lim_{N\to\infty} \mathbf{E}\,\langle Q_{2k}^{(N)}(p)\rangle = B_k(p); \tag{7}$$

the limit is given by equality

$$B_k(p) = p\phi_k(p^2),$$

where $\phi_k(p)$ are determined by recurrent relations

$$\phi_{k+1}(p) = \sum_{i=1}^{k+1} p^i\, \phi_{i-1}(p)\, \phi_{k+1-i}(p),$$

with initial conditions $\phi_0(p) = 1,\ \phi_1(p) = 1$. The numbers $C_k(q) = \phi_k(p)\, p^{\binom{k-1}{2}}$, where $q = p^{-1}$, are known as the q-Catalan numbers [7].

To discuss this result, let us note that the numbers $B_k(p)$ are represented by the sum over all possible partitions $\omega \in \hat{\Omega}_{2k}$ of the points $1, 2, \ldots, 2k$ to the set of k non-crossing pairs; each partition is weighted by powers of p (cf.(4))

$$B_k(p) = \sum_{\omega\in\hat{\Omega}_{2k}} \prod_{(i,j)\in\omega} p^{|i-j|}, \tag{8}$$

such that

$$t_k = \sum_{\omega\in\hat{\Omega}_{2k}} 1 = \frac{(2k)!}{k!(k+1)!}.$$

Each non-crossing partition into pairs $\omega \in \hat{\Omega}_{2k}$ can be identified with one of the half-plane rooted trees T_k of k edges [7]. Thus t_k represents the total number of the elements $|T_k|$.

Regarding definition (8), one can easily deduce that $B_k(p)$ satisfy the following recurrent relations

$$B_{k+1}(p) = \sum_{i=1}^{k+1} p^{2i-1} B_{i-1}(p) B_{k+1-i}(p) \tag{9}$$

with the conditions $B_0(p) = B_1(p) = 1$. Indeed, following the reasoning by Wigner [9], let us consider the subsum of (8) over those partitions, where the first point 1 is paired with the last point $2k$; we denote by $B_k'(p)$ the corresponding weighted sum. Then it is easy to observe that

$$B_k'(p) = p^{2k-1} B_{k-1}(p).$$

To complete the derivation of (9), it remains to consider the sum over partitions where 1 is paired with $2i$. Relations (8) and (9) answer the question of C. Mazza [3] and theorem 1.1 establishes relations of (8) with the product of random matrices of infinite dimensions.

In this connection, let us make one more remark about relations of our results with the non-commutative probability theory [8]. In frameworks of this approach, the set of random matrices $A^{(r)}$ with $V(r) = \delta_{0,r}$ represents in the limit $N \to \infty$ a family of free random variables $X^{(r)}$ with respect to the mathematical expectation $\mathbf{E}\langle\cdot\rangle$ (6). Free random variables represent a non-commutative analogue of jointly independent scalar random variables. In particular, according to the rules adopted to compute the moments of free random variables (see e.g. [6]), we recover the (ordinary) Catalan numbers

$$\lim_{N\to\infty} \langle [A_N^{(1)}]^{2k} \rangle = B_k(1) = t_k$$

known since the pioneering work by Wigner on the eigenvalue distribuion of large random matrices [9]. In [5] one can find a detailed analysis of the multiplicative functions over the non-crossing partitions, where, in particular, several generalizations of the Catalan numbers appear. However, the q-Catalan numbers are not present in [5].

In this context, conditions (1) can be regarded as a starting point to define the non-commutative analogs $Y^{(r)}$ of the correlated scalar (gaussian) random variables. The standard rules to compute the average value $\langle Y^{(1)} \cdots Y^{(2k)} \rangle$ could be added there by conditions

$$\langle Y^{(s)} Y^{(t)} \rangle = V(s-t).$$

Returning to the generalization of the Catalan numbers (8), let us note that as well as for the scalar case (5), there should be a critical value p'_c in the sense that the limit $N \to \infty$ of (5) exhibits different behaviour as $k \to \infty$ in dependence whether $p > p'_c$ or $p < p'_c$. This can be explained by the observation that the trees of T_k that have a vertex of large degree are relatively rare (see e.g. [2]). However, the weight (8) with $p < 1$ ascribes to such trees the probability greater than, say, to binary trees. This makes the subject of weighted non-crossing pairings reach and interesting.

Proof of Theorem 1.1.

Our goal is to derive recurrent relation (9). We rewrite (6) in the form

$$\langle Q_{2k}^{(N)}(p) \rangle = \frac{1}{N} \sum_{\{i\}} A^{(1)}_{i_1 i_2} A^{(2)}_{i_2 i_3} \cdots A^{(2k)}_{i_{2k} i_1} \equiv \langle Q^{(N)}_{(1,\ldots,2k)}(p) \rangle$$

and compute the mathematical expectation with the help of (3) and then (1):

$$\mathbf{E}\,\langle Q_{2k}^{(N)}(p)\rangle = \frac{1}{N^2}\sum_{\{i\}}\sum_{l=2}^{2k} V(l-1)\times$$

$$\mathbf{E}\,\left\{A^{(2)}_{i_2 i_3}\cdots A^{(l-1)}_{i_{l-1}i_l}\left(\delta_{i_1 i_l}\delta_{i_2 i_{l+1}}+\delta_{i_1 i_{l+1}}\delta_{i_2 i_l}\right)A^{(l+1)}_{i_{l+1}i_{l+2}}\cdots A^{(2k)}_{i_{2k}i_1}\right\}.$$

In this relations we mean that for $l=2$ and $l=2k$ the expression in curly brackets takes the forms $A^{(3)}\cdots A^{(2k)}$ and $A^{(2)}\cdots A^{(2k-2)}$, respectively.

Then we obtain relation

$$\mathbf{E}\,\langle Q_{2k}^{(N)}(p)\rangle = \sum_{l=2}^{2k} V(l-1)\mathbf{E}\,\langle Q^{(N)}_{(2,3,\ldots,l-1)}(p)\rangle\langle Q^{(N)}_{(l+1,\ldots,2k)}(p)\rangle +$$

$$\frac{1}{N}\sum_{l=1}^{2k} V(l-1)\mathbf{E}\,\langle Q^{(N)}_{(2,\ldots,l-1)}(p)Q^{(N)}_{(2k,\ldots,l+1)}(p)\rangle. \tag{10}$$

To complete the proof of Theorem 1.1, it remains to prove the following three items:

1) the moments of the normalized traces factorize

$$\mathbf{E}\,\langle Q^{(N)}_{(2,3,\ldots,l-1)}(p)\rangle\langle Q^{(N)}_{(l+1,\ldots,2k)}(p)\rangle - \mathbf{E}\,\langle Q^{(N)}_{(2,3,\ldots,l-1)}(p)\rangle\;\mathbf{E}\,\langle Q^{(N)}_{(l+1,\ldots,2k)}(p)\rangle = o(1) \tag{11}$$

in the limit $N\to\infty$;

2) the odd moments are zero

$$\mathbf{E}\,\langle Q^{(N)}_{(1,2,\ldots,2m+1)}(p)\rangle = 0 \tag{12}$$

and the average is invariant with respect to simultaneous shifts of all values of the subscripts

$$\mathbf{E}\,\langle Q^{(N)}_{(2m+1,2m+2,\ldots,2k)}(p)\rangle = \mathbf{E}\,\langle Q^{(N)}_{(1,2,\ldots,2k-2m)}(p)\rangle \tag{13}$$

and finally

3) the value of

$$\mathbf{E}\,\langle Q^{(N)}_{(2,\ldots,l-1)}(p)Q^{(N)}_{(2k,\ldots,l+1)}(p)\rangle \tag{14}$$

remains bounded as $N\to\infty$.

With these items in mind, we can easily derive from (10) that (7) takes place and $B_k(p)$ satisfy (9).

Relations (12) and (13) trivially follow from the definitions. Regarding (14), we observe that it is equal to

$$\mathbf{E}\,\langle Q^{(N)}_{(1,\ldots,l-2,2k-1\ldots,l)}(p)\rangle,$$

where the number of factors A is $2k-2$. Now the question of its behavior as $N \to \infty$ is reduced to the problem of computing the expected value of (6) where $2k-2$ factors are subjected to certain permutation. Thus, (14) can be estimated in terms of $\mathbf{E}\,\langle Q_{2k-2}(p)\rangle$ that is sufficient for us.

Let us describe the scheme of the proof of (11) based on the following standard procedure (see e.g. [1]). Let us denote $\xi^\circ = \xi - \mathbf{E}\,\xi$. Then we can write relations

$$R_2^{(N)}([l+1,\dots,2k],[2,3,\dots,l-1]) \equiv \mathbf{E}\,\langle Q^{(N)}_{(l+1,\dots,2k)}(p)\rangle^\circ \langle Q^{(N)}_{(2,3,\dots,l-1)}(p)\rangle =$$

$$\mathbf{E}\left\{ \langle Q^{(N)}_{(l+1,\dots,2k)}(p)\rangle^\circ \frac{1}{N}\sum_{\{i\}} A^{(2)}_{i_2 i_3}\cdots A^{(l-1)}_{i_{l-1}i_2}\right\}.$$

To compute the last mathematical expectation, we use again the identity (3) with $\gamma_l = A^{(2)}_{i_2 i_3}$. Repeating the same computations as above, we derive with the help of (1) equality

$$\mathbf{E}\,\langle Q^{(N)}_{(l+1,\dots,2k)}(p)\rangle^\circ \langle Q^{(N)}_{2,3,\dots,l-1}(p)\rangle =$$

$$\sum_{j=3}^{l-1} V(2-j)\mathbf{E}\left\{ \langle Q^{(N)}_{(l+1,\dots,2k)}(p)\rangle^\circ \langle Q^{(N)}_{(3,\dots,j-1)}(p)\rangle \langle Q^{(N)}_{(j+1,\dots,l-1)}(p)\rangle\right\} +$$

$$\frac{1}{N}\sum_{j=3}^{l-1} V(2-j)\mathbf{E}\left\{ \langle Q^{(N)}_{(l+1,\dots,2k)}(p)\rangle^\circ \langle Q^{(N)}_{(3,\dots,j-1)}(p) Q^{(N)}_{(j+1,\dots,l-1)}(p)\rangle\right\} +$$

$$\frac{1}{N^2}\sum_{j=l+2}^{2k} V(2-j)\mathbf{E}\left\{ \langle Q^{(N)}_{(2k,2k-1,\dots,j+1)}(p) Q^{(N)}_{(3,\dots,l-1)}(p) Q^{(N)}_{(l+1,\dots,j-1)}(p)\rangle\right\} +$$

$$\frac{1}{N^2}\sum_{j=l+2}^{2k} V(2-j)\mathbf{E}\left\{ \langle Q^{(N)}_{(2k,2k-1,\dots,j+1)}(p) Q^{(N)}_{(l-1,l-2,\dots,3s)}(p) Q^{(N)}_{(l+1,\dots,j-1)}(p)\rangle\right\}. \quad (15)$$

This somewhat cumbersome relation has rather simple structure. Indeed, using identity

$$\mathbf{E}\,\xi^\circ\zeta\mu = \mathbf{E}\,\xi^\circ\zeta\mathbf{E}\,\mu + \mathbf{E}\,\xi^\circ\mu\mathbf{E}\,\zeta + \mathbf{E}\,\xi^\circ\zeta^\circ\mu^\circ,$$

we observe that $R_2 = \mathbf{E}\,Q^\circ Q^\circ$ is expressed as the sum of terms of the following forms: $R_2\mathbf{E}\,Q$, $R_3 = \mathbf{E}\,Q^\circ Q^\circ Q^\circ$, R_2/N and $\mathbf{E}\,Q/N^2$. Also we see that in (15) the total number of factors A is decreased by 2.

Thus, we conclude that repeating this procedure with respect to all terms R_m in (15), we obtain the finite number (depending on k) of terms that involve products of V and $1/N$. The only expressions that has no factor $1/N$ are

$$R_m = \mathbf{E}\left\{\prod_{j=1}^m \langle Q^{(N)}_{(\alpha_j,\beta_j)}\rangle^\circ\right\}, m \le k/3.$$

This averaged product can be treated as before with the help of (3). We obtain relations that express R_m in terms of the sum of R_{m-1} and terms that have factors $1/N$. Taking into account that k is finite, we arrive at (11). Theorem 1.1 is proved. $\square$

References

[1] A. Boutet de Monvel and A. Khorunzhy, On the norm and eigenvalue distribution of large random matrices, *Ann. Probab.* **27** (1999) 913-944

[2] A. Khorunzhy, Sparse random matrices: spectral edge and statistics of rooted trees, *Adv. Appl. Probab.* **33** (2001) 124-140

[3] C. Mazza, *private communication*

[4] M L Mehta, *Random Matrices*, Acad. Press (1991)

[5] R. Speicher, Multiplicative functions on the lattice of non-crossing partitions and free convolution, *Math. Annalen* **298** (1994) 611-628

[6] P. Neu and R. Speicher, Rigorous mean field model for CPA: Anderson model with free random variables. *J. Stat. Phys.* **80** (1995) 1279-1308

[7] R.P. Stanley, *Enumerative Combinatorics*, vol.II, Cambridge University Press (1999)

[8] D.V. Voiculescu, K. Dykema and A. Nica, *Free Random Variables*, CRM Monograph Series, No. 1, Provodence, RI (1992)

[9] E. Wigner, Characteristic vectors of bordered matrices with infinite dimensions, *Ann. of Math.* **62** (1955) 548-564